Amendments to the Specification:

Please replace the title beginning at page 1, line 1 with the following replacement title:

--GENE [~~CLUSTER~~] ENCODING A NONRIBOSOMAL PEPTIDE SYNTHETASE FOR THE PRODUCTION OF RAMOPLANIN [~~BIOSYNTHESIS~~]--

Please replace the abstract paragraph with the following replacement paragraph:

-- The present invention relates to an isolated genetic sequence[s] encoding a nonribosomal peptide synthetase (NRPS) protein[s] which directs the biosynthesis of the antibiotic ramoplanin in microorganisms such as *Actinoplanes* sp. The isolated [~~biosynthetic~~] gene [~~cluster~~] sequence serves as a substrate for bioengineering of antibiotic structures.--

Please replace the paragraph [0012] beginning at page 4, line 1 with the following replacement paragraph:

-- In one embodiment preferred nucleic acids encode at least two, more preferably three, still more preferably four, or most preferably five or more ORFs selected from ORFS 1 to 33 (SEQ ID NOS: 2 to 34) of the ramoplanin locus. In one embodiment, combinations of ORFs selected from ORFs 1 through 33 (SEQ ID NOS 2 to 34) are provided which encode polypeptides that form at least the depsipeptide core structure of ramoplanin. In another embodiment combinations of ORFs selected from ORFs 1 through 33 (SEQ ID NOS: 2 to 34) are provided which encode polypeptides that form at least the fatty-aid side chain of the depsipeptide core structure of ramoplanin. In another embodiment, combinations of ORFs selected from ORFs 1 through 33 (SEQ ID NOS: 2 to 34) are provided which encode polypeptides responsible for the synthesis of 4-hydroxyphenylglycine (HPG) of ramoplanin. In another embodiment, combinations of ORFs selected from ORFs 1 through 33 (SEQ ID NOS: 2 to 34) are provided that encode polypeptides that form at least the beta-hydroxyasparagine residue. In another

embodiment, combinations of ORFs selected from ORFs 1 through 33 (SEQ ID NOS: 2 to 34) are provided which are involved in the regulation of ramoplanin biosynthesis. In another embodiment, combinations of ORFs selected from ORFs 1 through 33 (SEQ ID NOS: 2 to 34) are provided which encode polypeptides that are involved in resistance and subcellular localization of the ramoplanin biosynthetic machinery. A single ORF or a combination of ORFs selected from ORFs 1 through 33 (SEQ ID NOS: 2 to 34) are provided to enhance production of ramoplanin by altering the expression level of an ORF selected from ORFs 1 through 33 (SEQ ID NOS: 2 to 34). In another embodiment, the expression level of an ORF selected from ORFs 1 through 33 (SEQ ID NOS: 2 to 34) may be altered to increase the yield of a particular form of ramoplanin.--

Please replace the paragraph [0028] beginning at page 8, line 29 with the following replacement paragraph:

-- Figure 3A is a clustal analysis of adenylation domains of ramoplanin biosynthetic enzymes (amino acids 471-959 of SEQ ID NO:13 (ORF 12), amino acids 518-990 of SEQ ID NO:14 (ORF 13 M1), amino acids 1561-2052 of SEQ ID NO:14 (ORF 13 M2), amino acids 2619-3122 of SEQ ID NO:14 (ORF 13 M3), amino acids 3698-4160 of SEQ ID NO:14 (ORF 13 M4), amino acids 4719-5192 of SEQ ID NO:14 (ORF 13 M5), amino acids 6318-6804 of SEQ ID NO:14 (ORF 13 M7), amino acids 487-993 of SEQ ID NO:15 (ORF 14 M1), amino acids 1568-2041 of SEQ ID NO:15 (ORF 14 M2), amino acids 2603-3095 of SEQ ID NO:15 (ORF 14 M3), amino acids 3672-4135 of SEQ ID NO:15 (ORF 14 M4), amino acids 4699-5199 of SEQ ID NO:15 (ORF 14 M5), amino acids 5777-6280 of SEQ ID NO:15 (ORF 14 M6), amino acids 6840-7343 of SEQ ID NO:15 (ORF 14 M7), amino acids 7926-8380 of SEQ ID NO:15 (ORF 14 M8), and amino acids 309-804 of SEQ ID NO:18 (ORF 17), as defined in table 3). Shown is the alignment of the amino acid sequence (single letter code) of all adenylation domains found in the ramoplanin locus relative to the adenylation domain of gramicidin S synthetase GrsA (SEQ ID NO:35). Adenylation domains of multimodular non-ribosomal peptide synthetases ORF13 and ORF14 are labeled according to their corresponding module M1-M7 and M1-M8, respectively. Note that ORF13 does not contain an

adenylation domain in module 6. Highly conserved core motifs A1-A10 of adenylation domains (Konz et al., 1999, Chem. Biol. Vol. 6, pp. R39-48) are highlighted by boxes. Key residues used to predict the substrate specificity of each adenylation domain are highlighted in black (see Figure 3B).--

Please replace the paragraph [0030] beginning at page 9, line 24 with the following replacement paragraph:

-- Figure 3C shows the similarity between ORF26 (SEQ ID NO:27) and acyl-CoA ligases. Shown is the clustal analysis of ORF 26 versus several acyl-Coenzyme A ligases from diverse species: Mb, *Mycobacterium bovis* (SEQ ID NO:36); Mt, *Mycobacterium tuberculosis* (SEQ ID NO:37); Sv, *Streptomyces verticillus* (SEQ ID NO:38); Mx, *Myxococcus xanthus* (SEQ ID NO:39); Bs, *Bacillus subtilis* (SEQ ID NO:40). Highlighted by boxes are the highly conserved core motifs AL1-AL8 of acyl-CoA ligases [~~as described by Du et al., 2000~~].--

Please replace the paragraph [0032] beginning at page 10, line 15 with the following replacement paragraph:

-- Figure 5 illustrates two clustal alignments. Figure 5A shows the local amino acid sequence homology between ORF 10 (amino acids 263 to 318 of SEQ ID NO: 11) and a key motif found in pfam 00753 (SEQ ID NO:41) involved in coordinating two zinc molecules in the beta-lactamase superfamily. (For information regarding the Pfam Families Datebase, see Bateman *et al.* Nucleic Acids Research, 2000, Vol. 28, No. 1, 263-266). 1SML (SEQ ID NO:42) represents one member of this superfamily for which a crystal structure showing the intimate interaction between the zinc molecule and the highlighted residues is available (Ullah et al., J. Mol Biol., 1998 Nov 20; 284(1):125-36). Figure 5B shows the local amino acid sequence homology between ORF 10 (amino acids 405 to 452 of SEQ ID NO: 11) and a key motif found in pfam 00067 (SEQ ID NO:43) involved in coordinating an iron molecule in cytochrome P450 monooxygenases.--

Please replace the paragraph [0080] beginning at page 25, line 9 with the following replacement paragraph:

-- Methods of cloning and expressing large nucleic acids such as gene clusters, including NRPS-encoding gene clusters, in cells including *Streptomyces* are well known to those skilled in the art (*see, e.g.,* Stutzman-Engwall and Hutchinson (1989) *Proc. Ntl. Acad. Sci. USA,* 86: 3135-3139: Motamedi and Hutchinson (1987) *Proc. Natl. Acad. Sci. USA,* 84: 4445-4449; Grimm *et al.* (1994) *Gene,* 151 : 1-10; Kao *et al.* (1994) *Science,* 265 : 509-512; and Hopwood *et al.* (1987) *Meth. Enzymol.,* 153: 116-166). In some examples, nucleic acid sequences of well over 100 kb have been introduced into cells, including prokaryotic cells, using vector-based methods (*see* for example, Osoegawa *et al.,* (1998) *Genomics,* 52: 1-8; [~~Woon~~] Huang *et al.*, (1996) *Nucl. Acids, Res.*, 24: 4202-4209). --

Please replace the paragraph [0115] beginning at page 47, line 8 with the following replacement paragraph:

-- Yet another explanation is that specialized domains within the NRPSs may have evolved the ability to carry out dual functions. One domain that stands out as a candidate for having such dual functions is the condensation domain. Normally within a typical NRPS module that introduces a D-amino acid into the peptide product, epimerization (E) domains follow the thiolation (T) domain. In terms of linear domain organization on NRPS enzymes condensation (C) domains and epimerization (E) domains can be thought of occupying equivalent positions. That is, in an NRPS with multiple modules that is devoid of E domains, a C domain from any given module is found directly adjacent to the thiolation (T) domain of the upstream module. In addition, C domains and E domains also share a considerable amount of sequence similarity. Several highly conserved core motifs are shared between C and E domains. One particularly important motif that is common to both C and E domains is the histidine motif HHXXXDG (SEQ ID NO: 44) which has been shown by mutagenesis to form part

of the active site (Stachelhaus *et al.*; *Journal of Biological Chemistry* 1998;273:22773-22781). Thus, the C domains of modules 2, 3, 4 and 6 of OFR 13 (SEQ ID NO:14) and modules 2, 4 and 8 of ORF 14 (SEQ ID NO: 15) may be capable of amino acid epimerization as well as amide bond formation and be responsible for the 7-D-amino acid residues found in ramoplanin. --

Please replace the paragraph [0123] beginning at page 51, line 3 with the following replacement paragraph:

-- Figure 5 illustrates clustal alignments showing sequence homology between ORF [~~11~~] 10 (SEQ ID NO: [~~12~~] 11) and various metal ligand motifs. In each of the clustal alignments: (i) a line above the alignment is used to mark strongly conserved positions; (ii) an asterisk "*" indicates positions which have a single, fully conserved residues; (iii) a colon ":" indicates that one of the following strong groups is fully conserved: S, T or A; N, E, Q or K; N, H, Q or K; N, D, E or Q; Q, H, R or K; M, I, L or V; M, I, L or F; H or Y; and F, Y or W; and (iv) a period "." indicates that one of the following weaker groups is fully conserved: C, S or A; A, T or V; S, A or G; S, T, N or K; S, T, P or A; S, G, N or D; S, N, D, E, Q or K; N, D, E, Q, H or K; N, E, Q, H, R or K; F, V, L, I or M: and H, F or Y.--

Please replace the paragraph [0124] beginning at page 51, line 11 with the following replacement paragraph:

-- ORF 10 (SEQ ID NO: 11) contains two amino acid sequence motifs that are frequently found in enzymes that use metal cofactors. The N-terminal region of ORF 10 (SEQ ID NO: 11) contains a cluster of histidine residues (the His-motif) that shows significant local sequence homology to a conserved histidine motif found in several zinc-binding beta-lactamases. Figure 5A shows the local amino acid sequence homology between ORF 10 (SEQ ID NO: 11) and a key motif involved in coordinating two zinc molecules in the beta-lactamase superfamily. The alignment depicts amino acids 263 to 318 of ORF 10 (SEQ ID NO: 11), amino acids 42 to 99 of a member of the beta-lactamase superfamily, the L1 metallo-beta-lactamase (1SML) from *Stenotrophomonas*

maltophilia for which the crystal structure has been determined (Ullah *et al.,* 1998, J. Mol. Biol., 125-136), and amino acids 12 to 67 of the consensus sequence for pfam00753, i.e. the beta-lactamase superfamily motif (Bateman *et al.*, 2000, *Nucleic Acids Research*, Vol. 28, No. 1, 263-266). Highlighted in black are residues demonstrated in the L1 metallo-beta-lactamase to co-ordinate zinc and their counterparts in the other two sequences. X-ray crystal structure analysis demonstrates that the histidine residues in this conserved motif are responsible for binding the zinc metal cofactor (Ullah *et al.,* 1998, J. Mol. Biol., 125-136). The precise alignment and conserved spacing of the amino acid residues in the His-motif of ORF 10 (SEQ ID NO: 11) as compared to the zinc-binding beta-lactamases indicates that ORF 10 (SEQ ID NO: 11) is likely to bind a metal cofactor. --

Please replace the paragraph [0125] beginning at page 51, line 30 with the following replacement paragraph:

-- Figure 5B shows the local amino acid sequence homology between ORF 10 (SEQ ID NO: 11) and a key motif involved in coordinating an iron molecule in cytochrome P450 monooxygenases. The alignment depicts amino acids 405 to 452 of ORF 10 (SEQ ID NO: 11) and amino acids 370 to 421 of the consensus sequence for pfam00067, i.e. the cytochrome P450 motif (Bateman *et al.*, 2000, *Nucleic Acids Research*, Vol. 28, No. 1, 263-266). The region of ORF 10 (SEQ ID NO: 11) in highlight is in relatively good agreement with the Prosite motif PS00086 (Hofmann *et al.*, 1999, *Nucleic Acids Research*, Vol. 27, No. 1, 215-219) required for binding iron: [FW]-[SGNH]-x-[GD]-x-[RKHPT]-x-C-[LIVMFAP]-[GAD], where x is [~~any amino acid~~] a gap and amino acids in brackets indicate the variability in a given position. Notably, the least variable positions of this motif are present in ORF10 (SEQ ID NO: 11), i.e. residues Phe-423, Gly-425, Cys-428, and Gly-430). The C-terminal region of ORF 10 (SEQ ID NO: 11) contains a cluster of amino acid residues that shows significant local sequence homology to a motif frequently found in cytochrome P450 monooxygenases (the Cys-motif). This motif includes a cysteine residue that is highly conserved in the cytochrome P450 monooxygenases and that has been shown by X-ray crystal structure analysis to be

involved in binding the iron metal cofactor required for catalysis. The Cys-motif of ORF 10 (SEQ ID NO: 11) is likely to contribute to the binding of a metal cofactor. The presence of two amino acid sequence motifs that are found in well-characterized metal-binding enzymes indicates that ORF 10 (SEQ ID NO: 11) is likely to be a metal-binding enzyme. Thus, the ORF 10 (SEQ ID NO: 11) is likely to be responsible for the formation of beta-hydroxyasparagine during the synthesis of ramoplanin.--

Please replace the paragraph [0130] beginning at page 53, line 28 with the following replacement paragraph:

-- As an alternative source of native ORF 10 (SEQ ID NO: 11), a *Streptomyces* expression system was employed. ORF 10 (SEQ ID NO: 11) was amplified by high fidelity PCR using two specific oligonucleotides, namely primer sequences (5' to 3') N-oligo: CACACAGAATTCACCAGCGCCACTCGCGCTT (SEQ ID NO:45), and C-oligo: CACACATCGATGGGCAACGCCGATCAGCCG (SEQ ID NO:46). This primer pair introduces convenient restriction enzyme sites at either end of the ORF 10 gene but does not introduce any exogenous amino acids. The amplified genes were then subcloned using *Cla*I and *Eco*RI restriction enzymes into a *Streptomyces*/E.coli expression shuttle vector, pECO1202. Following confirmation of the cloned sequences, *Streptomyces lividans* TK24 was transformed with this construct. Five independent transformants were selected for further analysis. Cultures were grown for 48 hours in a gyrating 30°C incubator using 25 ml erlenmeyer flasks containing 5 ml of Tryptic Soy Broth (TSB, Difco). Total RNA was extracted from the cell pellets using the RNeasy kit (Qiagen). The integrity and concentration of the RNA was monitored by agarose gel electrophoresis. Subsequently, reverse transcription was performed using 1 ug total RNA primed with an antisense primer sequence located in the vector just downstream of the stop codon. Following reverse transcription of each sample and appropriate controls, 20 cycles of PCR were performed using the original ORF-specific oligonucleotides, N-oligo and C-oligo. According to the RT-PCR analysis, the five recombinant *S. lividans* clones express relatively high levels of ORF 10-specific mRNA and the size of the RT-PCR product is as expected. Figure 6 shows the RT-PCR

analysis of recombinant *S. lividans* clones expressing ramoplanin ORF 10, wherein is lane 1 is 1 kb DNA ladder; lane 2 is untransformed *S. lividans*; lane 3 is *S. lividans* transformed with empty expression vector; lanes 4-8 are five different *S. lividans* recombinant clones expressing ramoplanin orf 10; lane 9 is an *S. lividans* recombinant clone expressing an unrelated gene; lane 10 is negative control performed without RNA; lane 11 is negative control performed without RT; lane 12 is positive control for PCR using plasmid DNA.--

Please replace Table 2 on pages 35 -41 with the following replacement Table:

ORF	SEQ ID NO	# aa	proposed function	GenBank accession	probability	% identity	% similarity	proposed function of GenBank match
1	2	333	unknown; membrane protein	CAB48902	5.00E-22	27	41	possible membrane protein, unknown function, in Streptomyces coelicolor
2	3	304	ABC transporter	CAB48901	3.00E-55	42	59	probable ABC transporter ATP-binding protein from Streptomyces coelicolor
				AAF81232	7.00E-32	31	47	ABC transporter ATP binding protein found in nonactin biosynthetic locus of Streptomyces griseus
				AAF12291	4.00E-29	34	51	ABC transporter, ATP-binding protein from Deinococcus radiodurans
3	4	[~~321~~] 336	unknown; membrane protein	CAB48902	2.00E-15	35	50	possible membrane protein, unknown function, in Streptomyces coelicolor
4	5	283	oxidoreductase similar to prephenate dehydrogenases	CAA11792	2.00E-69	53	63	similar to prephenate dehydrogenase; chloroeremomycin biosynthesis in Amycolatopsis orientalis
				CAB38592	2.00E-67	50	62	probable oxidoreductase similar to prephenate dehydrogenase; calcium-dependent antibiotic biosynthesis in Streptomyces coelicolor
				AAF67499	3.00E-66	47	64	putative oxidoreductase protein similar to prephenate dehydrogenase; novobiocin biosynthesis in Streptomyces spheroides
5	6	336	transcriptional regulator similar to StrR	CAA07385	1.00E-74	46	58	StrR DNA-binding protein/regulator of 5'-hydroxystreptomycin biosynthesis in Streptomyces glaucescens; positive transcriptional regulator of strU, strVW genes
				CAB45047	2.00E-74	47	62	probable transcriptional regulator in chloroeremomycin biosynthetic locus of Amycolatopsis orientalis; similar to other regulators of antibiotic biosynthesis
				CAA68515	4.00E-70	47	60	putative regulatory protein StrR in streptomycin biosynthetic locus in Streptomyces griseus
				AAB66654	6.00E-68	44	59	SpcR putative transcriptional regulator of spectinomycin biosynthesis in Streptomyces flavopersicus
				AAF67500	9.00E-58	42	61	NovG putative regulatory protein in novobiocin biosynthetic locus of Streptomyces spheroides

ORF	SEQ ID NO	# aa	proposed function	GenBank accession	probability	% identity	% similarity	proposed function of GenBank match
6	7	444	Amino-transferase	CAB38598	1.00E-123	56	67	possible aminotransferase found in the calcium-dependent antibiotic biosynthetic locus of Streptomyces coelicolor
				CAA11790	1.00E-101	47	62	protein similar to aminotransferase found in the chloroeremomycin biosynthetic locus of Amycolatopsis orientalis
7	8	356	oxidoreductase similar to glycolate oxidases	CAB38520	1.00E-115	60	70	putative glycolate oxidase found in calcium-dependent antibiotic biosynthetic locus of Streptomyces coelicolor
				AAA34030	6.00E-77	47	62	spinach glycolate oxidase from Spinacia oleracea
				CAB78838	2.00E-75	45	60	glycolate oxidase-like protein from Arabidopsis thaliana
				CAA11762	4.00E-75	47	61	protein similar to glycolate oxidase in chloroeremomycin biosynthetic locus of Amycolatopsis orientalis
8	9	640	ABC transporter involved in resistance/transport	CAA11793	0	55	71	protein similar to mdr/ABC transporter found in chloroeremomycin biosynthetic locus of Amycolatopsis orientalis
				AAF67494	1.00E-114	38	57	NovA ABC transporter in novobiocin biosynthetic locus of Streptomyces spheroides
				CAB38879	1.00E-78	34	50	probable ABC transporter found in the calcium-dependent antibiotic biosynthetic locus of Streptomyces coelicolor
9	10	271	esterase/hydrolase	CAB38877	6.00E-66	48	63	probable hydrolase found in the calcium-dependent antibiotic biosynthetic locus of Streptomyces coelicolor
				CAA11784	9.00E-58	44	56	protein similar to haloperoxidase found in chloroeremomycin biosynthetic locus of Amycolatopsis orientalis
				CAA71338	2.00E-45	41	54	putative thioesterase found in streptothricin biosynthetic locus of Streptomyces sp. strain F20
10	11	529	unknown	AAB30311	2.00E-29	41	56	unknown protein found in putative chloramphenicol biosynthetic locus of Streptomyces venezuelae
11	12	90	acyl carrier protein	AAA22001	6.00E-08	33	54	polyketide synthase in Anabaena PCC7120
				CAA98988	8.00E-08	37	57	polyketide synthase found in the phenolpthiocerol biosynthetic locus of Mycobacterium tuberculosis

ORF	SEQ ID NO	# aa	proposed function	GenBank accession	probability	% identity	% similarity	proposed function of GenBank match
				AAF62883	7.00E-07	39	55	type I polyketide synthase found in the epothilone biosynthetic locus of Sorangium cellulosum
12	13	1051	nonribosomal peptide synthetase	CAB15186	0	38	55	nonribosomal peptide synthetase involved in siderophore 2,3-dihydroxybenzoate biosynthesis in Bacillus subtilis
				AAD56240	0	38	55	DhbF peptide synthetase involved in siderophore production in Bacillus subtilis
				AAC38442	1.00E-179	40	52	actinomycin synthetase II peptide synthetase found in the actinomycin biosynthetic locus of Streptomyces chrysomallus
13	14	6893	nonribosomal peptide synthetase	AAC80285	0	36	52	SyrE peptide synthetase found in the syringomycin biosynthetic locus of Pseudomonas syringae
				AAC45930	0	31	48	TycC tyrocidine synthetase 3 found in the tyrocidine biosynthetic locus of Brevibacillus brevis
14	15	8695	nonribosomal peptide synthetase	AAC80285	0	36	51	SyrE peptide synthetase found in the syringomycin biosynthetic locus of Pseudomonas syringae
				AAC45930	0	32	49	TycC tyrocidine synthetase 3 found in the tyrocidine biosynthetic locus of Brevibacillus brevis
15	16	234	thioesterase	AAC69333	2.00E-30	36	50	PikAV thioesterase II found in the methymycin/pikromycin biosynthetic locus of Streptomyces venezuelae
				AAC01736	6.00E-30	34	49	thioesterase found in the rifamycin biosynthetic locus of Amycolatopsis mediterranei
				CAA57967	2.00E-29	39	48	protein with similarity to thioesterases found in the pyochelin biosynthetic locus of Pseudomonas aeruginosa
				AAA79279	1.00E-28	34	48	thioesterase found in the bialaphos biosynthetic locus of Streptomyces hygroscopicus
16	17	274	short chain secondary alcohol dehydrogenase/	CAB54559	7.00E-49	39	58	Rhodococcus erythropolis LimC carveol dehydrogenase, a nicotinoprotein belonging to the short chain alcohol dehydrogenase/reductase superfamily
			3-ketoacyl-acyl carrier protein reductase	CAA15546	3.00E-46	39	54	hypothetical protein from Mycobacterium tuberculosis, similar to dehydrogenases
				AAF64503	9.00E-43	39	53	cholesterol oxidase from Nocardioides simplex
				CAA68181	2.00E-38	38	54	UcpA protein, belongs to alcohol dehydrogenase /rybitol dehydrogenase family

ORF	SEQ ID NO	# aa	proposed function	GenBank accession	probability	% identity	% similarity	proposed function of GenBank match
				AAC44307	4.00E-36	34	53	FabG 3-ketoacyl-acyl carrier protein reductase from Bacillus subtilis
				CAA77599	1.00E-33	36	49	beta ketoacyl reductase in unknown polyketide biosynthetic locus of Streptomyces cinnamonensis
17	18	891	threonine-specific adenylate ligase	CAA67248	1.00E-143	49	58	Pristinamycin I synthase 2 nonribosomal peptide synthetase in the pristinamycin biosynthetic locus of Streptomyces pristinaespiralis
				AAC38442	1.00E-141	49	57	actinomycin synthetase II nonribosomal peptide synthetase in the actinomycin biosynthetic locus of Streptomyces chrysomallus
				CAB38518	1.00E-138	48	58	CDA peptide synthetase I found in the calcium-dependent antibiotic biosynthetic locus of Streptomyces coelicolor
18	19	187	unknown	none				
19	20	415	transmembrane protein	CAB42730	2.00E-82	43	57	probable transmembrane protein from Streptomyces coelicolor
				CAB02537	5.00E-59	39	50	probable membrane protein from Mycobacterium tuberculosis
				AAF25828	2.00E-56	35	48	putative transmembrane protein Mycobacterium smegmatis
20	21	491	halogenase/hydroxylase	CAA11780	1.00E-180	63	76	protein similar to non-heme oxygenase/halogenase found in chloroeremomycin biosynthetic locus of Amycolatopsis orientalis
				CAA76550	1.00E-178	63	75	BhaA protein similar to halogenase, found in the balhimycin biosynthetic locus of Amycolatopsis mediterranei
				AAB49297	1.00E-176	62	74	hypothetical hydroxylase a found in the vancomycin biosynthetic locus of Amycolatopsis orientalis
				AAD24884	6.00E-37	30	46	PltA putative halogenase found in the pyoluteorin biosynthetic locus of Pseudomonas fluorescens
21	22	217	two-component response regulator	CAB59507	9.00E-58	52	71	Streptomyces coelicolor protein highly similar to various putative two-component response regulators
				CAA22374	8.00E-52	52	66	probable luxR family response regulator from Streptomyces coelicolor
				CAB50960	3.00E-51	49	66	probable two-component system response regulator from Streptomyces coelicolor

ORF	SEQ ID NO	# aa	proposed function	GenBank accession	probability	% identity	% similarity	proposed function of GenBank match
				CAB42025	3.00E-48	49	64	probable two-component system regulator from Streptomyces coelicolor
				CAB38597	3.00E-38	44	58	AbsA2, two component response regulator from Streptomyces coelicolor, acts as part of a two component signal transduction system
22	23	403	two-component sensory protein kinase	CAB42041	1.00E-38	37	48	probable two-component system sensor kinase from Streptomyces coelicolor
				CAB51250	1.00E-34	32	44	probable two-component system sensor kinase from Streptomyces coelicolor
				CAB89761	1.00E-34	34	42	probable two-component system sensor kinase from Streptomyces coelicolor
				CAB38596	3.00E-27	31	43	AbsA1, two component sensor kinase from Streptomyces coelicolor, acts as part of a two component signal transduction system
23	24	309	ABC transporter involved in resistance/transport	CAB48901	2.00E-45	41	55	probable ABC transporter ATP-binding protein from Streptomyces coelicolor
				CAB49966	4.00E-28	33	55	ATP-binding transport protein from Pyrococcus abyssi
				AAF12291	9.00E-28	38	56	ABC transporter, ATP-binding protein from Deinococcus radiodurans
24	25	553	acyl-CoA dehydrogenase	AAD45605	2.00E-18	25	44	isovaleryl-CoA dehydrogenase from Arabidopsis thaliana
				CAB55554	7.00E-18	24	43	isovaleryl-CoA dehydrogenase from Pisum sativum
				CAB46799	4.00E-16	29	44	probable acyl-CoA dehydrogenase from Streptomyces coelicolor
				CAA16488	9.00E-14	29	39	RedW acyl-coa dehydrogenase in the undecylprodigiosin biosynthetic locus of Streptomyces coelicolor
				AAF08800	3.00E-13	23	44	YngJ protein found in the mycosubtilin biosynthetic locus of Bacillus subtilis
25	26	585	acyl-CoA dehydrogenase	CAB61531	2.00E-27	26	43	FadE fatty acid acyl-CoA dehydrogenase found in Streptomyces lividans
				CAB07077	6.00E-22	24	39	Mycobacterium tuberculosis protein highly similar to acyl-CoA dehydrogenase
				CAA17679	2.00E-21	26	43	probable Acyl-CoA dehydrogenase found in Mycobacterium tuberculosis

ORF	SEQ ID NO	# aa	proposed function	GenBank accession	probability	% identity	% similarity	proposed function of GenBank match
26	27	587	acyl-CoA ligase	AAG02359	1.00E-115	45	56	BlmVI peptide synthetase in bleomycin biosynthetic locus of Streptomyces verticillus
				AAC44128	1.00E-94	38	53	Mx1 peptide synthetase B in saframycin biosynthetic locus of Myxococcus xanthus
				CAA16183	1.00E-85	37	49	polyketide synthase found in the undecylprodigiosin biosynthetic locus of Streptomyces coelicolor
				CAB05426	3.00E-84	35	51	Fad29 probable acyl-CoA synthetase found in Mycobacterium tuberculosis
				CAA17589	2.00E-82	36	51	Fad24 probable acyl-CoA synthetase found in Mycobacterium tuberculosis
				CAB01395	1.00E-81	35	50	Fad25 probable acyl-CoA synthetase found in Mycobacterium tuberculosis
				AAB52538	2.00E-78	34	50	acyl-CoA synthase from Mycobacterium bovis
				CAB36629	4.00E-78	35	52	putative acyl-CoA synthase from Mycobacterium leprae
27	28	75	unknown	CAB38589	1.00E-24	70	80	small conserved hypothetical protein found in the calcium-dependent antibiotic biosynthetic locus of Streptomyces coelicolor
				CAB08480	3.00E-22	67	77	MbtH possibly involved in mycobactin synthesis in Mycobacterium tuberculosis
				CAA11799	3.00E-19	74	89	hypothetical protein found in chloroeremomycin biosynthetic locus of Amycolatopsis orientalis
28	29	94	chorismate mutase-like protein	CAB02002	2.00E-15	50	69	hypothetical protein in Mycobacterium tuberculosis
				CAB82023	2.00E-11	46	59	hypothetical protein in Streptomyces coelicolor
				CAB72783	7.00E-03	36	59	chorismate mutase\prephenate dehydratase from Campylobacter jejuni
				AAC75649	6.00E-02	30	50	chorismate mutase-T and prephenate dehydrogenase protein from E. coli
29	30	619	membrane protein	CAB16086	2.00E-56	28	43	unknown protein in Bacillus subtilis
				CAA05568	4.00E-34	35	54	YkcB unknown protein in Bacillus subtilis
				CAB76994	0.01	26	35	putative integral membrane protein in Streptomyces coelicolor
				AAC18892	0.049	29	37	transmembrane protein from Streptomyces aureofaciens

ORF	SEQ ID NO	# aa	proposed function	GenBank accession	probability	% identity	% similarity	proposed function of GenBank match
30	31	355	4-hydroxyphenylpyruvate dioxygenase	CAA11761	5.00E-87	50	63	protein similar to hydroxyphenyl pyruvate dioxygenase found in the chloroeremomycin biosynthetic locus of Amycolatopsis orientalis
				CAB38519	1.00E-69	44	54	probable 4-hydroxyphenylpyruvic acid dioxygenase found in the calcium-dependent antibiotic biosynthetic locus of Streptomyces coelicolor
				CAB51008	2.00E-49	36	51	probable 4-hydroxyphenylpyruvic acid dioxygenase found in Streptomyces coelicolor
				AAA50231	3.00E-49	36	50	4-hydroxyphenylpyruvate acid dioxygenase from Streptomyces avermitilis
31	32	429	transmembrane transporter	CAB45049	4.00E-81	46	64	putative integral membrane ion antiporter found in the chloroeremomycin biosynthetic locus of Amycolatopsis orientalis
				BAA16991	3.00E-72	39	56	sodium/proton antiporter from Synechocystis sp.
				CAA23036	8.00E-65	37	57	putative sodium/protein exchanging protein from Arabidopsis thaliana
				AAF26906	1.00E-41	30	48	protein similar to sodium/proton and drug/proton antiporters found in the epothilone biosynthetic locus of Sorangium cellulosum
32	33	189	Unknown	CAB72201	1.00E-11	31	41	hypothetical protein in Streptomyces coelicolor
				CAB56690	2.00E-08	31	42	hypothetical protein in Streptomyces coelicolor
33	34	309	Unknown, incomplete	none				